基于混合损失联合调优与多尺度分类相结合的肺结节检测算法

姚宇瑾，张 利

(清华大学 电子工程系，北京 100084)

摘 要：针对 CT 图像的肺结节自动检测任务中检测灵敏度低及存在大量假阳性的问题，提出了一个基于混合损失的三维全卷积网络与基于注意力的多尺度三维残差网络相结合的肺结节检测方法。首先，基于相似度损失预训练三维全卷积网络，利用该网络筛选难例样本，并基于混合损失将难例与正样本进行联合调优得到候选结节检测网络，用于快速筛选疑似结节；然后，利用基于注意力的多尺度三维残差卷积网络对疑似结节进行分类，从候选结节中精确地分辨出真正结节。在 LUN16 数据集上，候选结节检测阶段的灵敏度在每个病例的假阳数目为 59.1 时达到 97.18%，检测系统的平均灵敏度为 0.880，表明本算法可以提高肺结节检测的灵敏度并有效控制假阳性，在 LUNA16 数据集上获得了更优的性能。

关键词：肺结节检测；混合损失；联合调优；注意力；多尺度

中图分类号：TP391.41 doi: 10.3969/j.issn.1001-3695.2018.03.0153

Pulmonary nodule detection via hybrid loss based joint fine-tuning and multi-scale classification

Yao Yujin, Zhang Li

(*Dept. of Electronic Engineering*, *Tsinghua University*, *Beijing* 100084, *China*)

Abstract: To solve the problem that lung nodule automatic detection methods for CT images can only get low sensitivities with a lot of false positives, this paper proposed an integrated method with hybrid loss based 3D fully convolutional network as candidate detection and attention-based multi-scale residual network as nodule classification. Firstly, this paper established a 3D fully convolutional network based on dice coefficient loss, and the network filtered hard negative samples uniting with positives to fine-tune. Then, it designed an attention based multi-scale 3D residual convolutional network to classify the candidate and recognize true nodules. Experiment results on the LUNA16 dataset show that the proposed method achieves the sensitivity of 97.18% at 59.1 false positives per scan in the candidate detection stage, and the whole system achieves the average sensitivity of 0.880, which demonstrates our proposed method can improve sensitivity with fewer false positives and achieve superior performance.

Key words: lung nodule detection; hybrid loss; joint fine-tune; attention; multi-scale

0 引言

肺癌是目前全球范围内致死率最高的癌症之一[错误!未找到引用源。]，及早发现并治疗能大大提高肺癌病人的存活时间。早期肺癌常表现为无症状的肺结节，因此肺结节的检测对早期肺癌诊断与治疗有着重要意义。低剂量胸部 CT 是目前肺癌普查最常用的诊断方法，很多微小的肺结节也能通过 CT 被检测到。然而人工筛查肺结节的工作量大且诊断难度大，尤其是随着早期肺癌普查的需求增加，需要花费大量人力与时间进行 CT 分析与诊断。因此，精确和高效的肺结节自动检测是非常有意义的。由于肺结节在尺寸、形状、密度分布上的差异巨大，所处的肺部环境多样，可能与血管、肺壁粘连或者独立，所以 CT 图像的肺结节自动检测充满挑战性。

典型的肺结节自动检测系统由两阶段组成：a）候选结节检测，从 CT 中筛选出疑似肺结节，包含大量假阳性结节；b）假阳性去除，通过更精细的分类算法去除假阳性结节并得到最终检测结果。传统的检测算法[2][错误!未找到引用源。]大多基于肺结节的形状和 CT 值特点采用图像处理方法获得候选结节，然后提取候选结节的局部特征并构建分类器进行分类。Murphy 等人[错误!未找到引用源。]设计了形状指数与曲率特征选取种子体素，再利用滞后阈值法进行种子生长从而得到候选结节，然后采用了由粗到细的两个 K-最近邻分类器[错误!未找到引用源。]去除假阳性结节。粗分类的描

收稿日期：2018-03-06；**修回日期**：2018-04-17

作者简介：姚宇瑾（1993-），女，浙江绍兴人，硕士研究生，主要研究方向为医学图像处理、模式识别、深度学习（yaoyj15@mails.tsinghua.edu.cn）；张利（1965-），男，教授，博导，博士，主要研究方向为图像处理、模式识别.

述子由候选结节体素块的几何特征与灰度特征共 18 个特征串联而成，细分类的描述子包含更丰富的形状、梯度特征等 135 个特征。Jacobs 等人[错误!未找到引用源。]采用双阈值法基于 CT 值筛选候选结节体素，并通过形态学操作去除部分容积效应，再利用连通域分析进行聚类得到候选结节，然后对候选结节邻域体素块提取灰度、纹理、形状和环境上下文特征，同时也采用了由粗到细的二阶段分类，分类器选取了集成的 GentleBoost[错误!未找到引用源。]。Retico 等人[错误!未找到引用源。]根据 Hessian 矩阵特征值与局部几何形状的关系设计滤波函数进行多尺度结节增强，通过检测局部极大值获得候选结节，然后计算候选结节三维邻域块内体素的梯度矩阵和 Hessian 矩阵，以它们的特征值作为描述子，并训练一个三层的神经网络作为分类器。

这些传统方法依赖于描述子的准确性，而肺结节的表现形式复杂多样，人工设计的描述子往往只能表现表层特征，很难挖掘有鉴别力的特征，因此很难区分真结节和类似结节（如血管、气管等）。近年来随着深度学习的快速发展，卷积神经网络逐渐被应用于医学图像处理的各个领域，它可以在数据驱动下自动提取有鉴别力的特征。基于卷积神经网络的研究[错误!未找到引用源。错误!未找到引用源。]思路一般为首先训练全卷积网络，从而直接得到每个断层的预测概率图，并设定阈值选出候选结节，再以候选结节质心为中心截取固定尺寸二维块，由一个卷积神经网络对其进行二分类从而去除所有假阳结节。由于 CT 本质上是三维体数据，仅利用横截面的二维信息是不够的，一些研究提出了基于多视角或三维卷积的算法以更好地利用体数据[错误!未找到引用源。错误!未找到引用源。错误!未找到引用源。]。Setio 等人[错误!未找到引用源。]对候选结节所在三维块截取了 9 个不同的二维截面，分别输入到含 9 个平行独立分支的分类网络中，网络的末端将 9 个分支的分类结果进行融合得到最终三维块类别。Dou 等人[错误!未找到引用源。]为了解决肺结节尺度差异大的问题，以候选结节质心为中心分别提取了三个不同尺度的三维块，并分别输入到三个不同的分类网络得到每个尺度下的类别概率，融合后得到最终分类结果。此外，也有一些研究提出了基于 Faster-RCNN[错误!未找到引用源。]的候选结节检测方法[错误!未找到引用源。错误!未找到引用源。]。Ding 等人[错误!未找到引用源。]针对结节尺寸相对于肺部占比小的特点，在最后一个卷积层后增加了反卷积层以恢复更精细的特征，同时设计了若干固定尺寸的 anchor 与结节尺寸相匹配。然而仅利用传统分割与分类的方法并不足以解决结节检测问题，结合肺结节的形态与分布特点，可以探索更多新的策略并从三维块中挖掘出更具鉴别力的信息从而进一步提升系统的性能。

本文提出了一个改进的肺结节检测框架，该框架对系统的两个阶段都进行了优化。首先候选结节检测阶段改进为一个基于混合损失进行联合调优的三维全卷积网络，先基于 Dice 损失[错误!未找到引用源。]预训练网络并由该网络筛选难例样本，再将难例与正样本基于混合损失进行联合调优，增强了对假阳性样本的鉴别能力。结节分类阶段设计了基于注意力的多尺度残差网络，利用候选结节信息提取局部注意区域并通过多尺度网络结构增强对多尺度结节的鉴别能力，从而提高分类准确率。本文的算法在大规模数据集 LUNA16[错误!未找到引用源。]上进行了验证，获得了显著的性能提升。此外，也通过对比实验验证了每一个改进算法的有效性。

1　方法

在本章提出了一个肺结节自动检测系统，系统框架如图 1 所示，包括改进的基于 3D U-net[错误!未找到引用源。]的候选结节检测和基于注意力的多尺度结节分类。

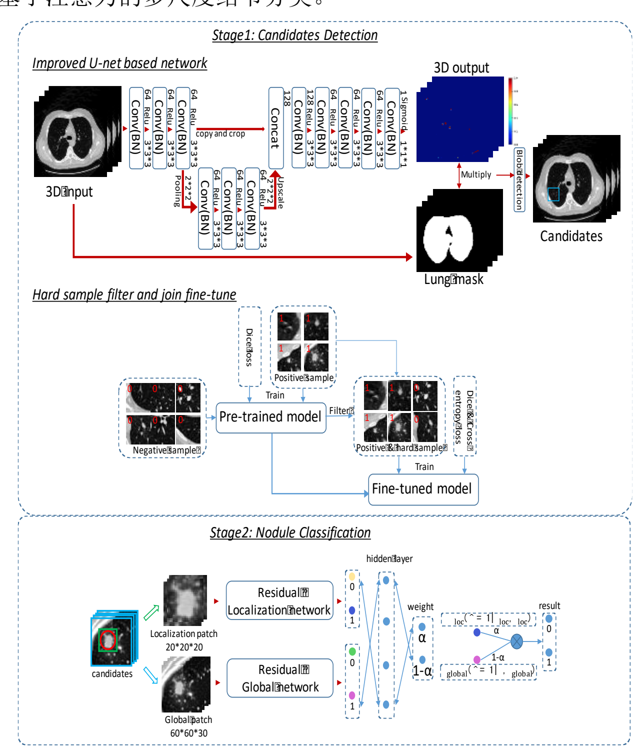


图 1　改进的肺结节检测系统框架

1.1　改进的基于 3D U-net 的候选结节检测

1.1.1 基于 Dice 损失的预训练模型

候选结节检测步骤首先对 CT 进行预筛选，选出疑似结节，并过滤掉其余绝大部分体素，从而减少输入到下一步分类网络中的样本数目。预筛选要保证结节检测的灵敏度，能够尽可能多地保留结节并滤除非结节。对此，本文设计了一个基于 3D U-net[错误!未找到引用源。]的分割网络。一方面，仅从二维图像区分血管和结节几乎是不可能的，因此网络中的所有卷积和池化操作都是三维的，从而充分利用 CT 体数据的空间信息；另一方面，全卷积网络将输出每个体素属于结节的概率分布情况，从而可以快速得到整个 CT 体数据的结节属性概率分布。如果以整个 CT 体数据作为训练输入会带来过大的计算量，而且判断任一体素 V_i是否属于结节由它的局部 CT 值分布决定，距离越远的体素对于V_i的预测结果影响越小，因此本文截取固定尺寸的三维块作为训练网络输入，在保证检测灵敏度的同时降低计算复杂度。由于输入三维块尺寸较小，所以对 U-net 进行了简化。具体网络结构见图 1。网络的输出是分辨率降低的三维体预测概率，与输入三维块存在对应关系。由于结节的平均直径不足 10 mm，对于包含结节的 56*56*56 的三维块而言，如果将每个体素看

做一个样本，则正负样本比约为 1:200。为了解决这种正负样本极其不平衡的情况，本文选择了改进后的 Dice 系数错误!未找到引用源。作为损失函数，从而避免模型向负样本倾斜。Dice 系数通过两个集合的重合率衡量相似情况，对于给定三维块{*X, Y, label*}，*X* 表示网络输入，*Y* 表示与 *X* 对应的标注，*label* 表示该三维块是否包含结节，1 是正样本包含结节，0 是负样本不包含结节，可得 Dice 损失函数为

$$L_{\text{dice}} = -\frac{2\times\sum_{\{\text{voxel }1,\dots,i,\dots,N\}} y_i * p(\hat{y}_i=1|\boldsymbol{X},\boldsymbol{W})}{\sum_{\{\text{voxel }1,\dots,i,\dots,N\}} y_i + \sum_{\{\text{voxel }1,\dots,i,\dots,N\}} p(\hat{y}_i=1|\boldsymbol{X},\boldsymbol{W})} \quad (1)$$

其中：$\boldsymbol{W}$ 表示网络参数；y_i表示每个体素的真实类别 0 或 1；$\hat{y}_i$表示每个体素的预测类别。对于包含结节的三维块，Dice 损失函数使得结节更倾向于被检测到，可以提高检测的灵敏度；但是对于不包含结节的三维块，Dice 损失函数恒为零，因此训练该网络的所有输入样本至少包含一个结节。

1.1.2 难例筛选与联合调优机制

上述算法可以识别一部分非结节，但缺乏对类似结节的负样本的鉴别能力。对于非结节，大部分是肺部背景区域，较易与结节区分，还有一部分则是血管、气管，它们的 CT 值与结节接近且局部形态相似，这类非结节虽然占少数但鉴别难度大，在训练中增加这类样本更有利于提高检测网络的鉴别能力。为了进一步利用这些难例样本，本文设计了难例挖掘错误!未找到引用源。并联合调优的机制。首先利用 1.1.1 节训练得到的模型对所有不含结节的负样本进行检测，预测概率大于设定阈值的体素被认为是假阳结节，所有包含假阳结节的负样本就是挖掘出来的难例。难例与正样本共同进行调优训练，但由于负样本的 Dice 损失恒为零，所以选择改进的交叉熵作为负样本的损失函数，正样本仍以 Dice 系数为损失函数，从而得到调优模型的损失函数为

$$L_{\text{ce}} = -\frac{1}{N}\sum_{\{\text{voxel }1,\dots,i,\dots,N\}} y_i \log p(\hat{y}_i=1|\boldsymbol{X},\boldsymbol{W}) + (1-y_i)\log(1-p(\hat{y}_i=1|\boldsymbol{X},\boldsymbol{W})) \quad (2)$$

$$L_{\text{neg}} = 2*\text{sigmoid}(L_{\text{ce}}) - 2,\ L_{\text{pos}} = L_{\text{dice}} \quad (3)$$

$$L = 1(label=1)\times L_{\text{pos}} + 1(label=0)\times L_{\text{neg}} \quad (4)$$

其中：L_{ce}是交叉熵损失的定义；$\text{sigmoid}(x) = 1/(1+e^{-x})$,对交叉熵损失进行 sigmoid 变换是为了使正负样本的损失值处于相同的范围；L_{pos}和L_{neg}分别为正负样本的损失；$1(label=1)$是指示函数，表示对正负样本采用不同的损失；L 为调优网络的混合损失函数。

在测试阶段，由于全卷积网络对输入尺寸没有限制，直接将整个 CT 体数据输入到网络中，得到体预测概率，而输出每个体素可以直接定位到对应的输入体素，从而可得整个 CT 的结节属性概率分布。通过对体预测概率进行极值检测即可获得 CT 中的所有疑似结节。

由于以上结节检测的训练网络的输入只包含局部信息，无法获得全局上下文信息，所以肺实质外部会产生大量假阳性结节。为了解决这一问题，本文采用了肺分割算法**错误!未找到引用源。**以滤除肺实质之外的假阳性结节，这也是许多基于肺的医学图像问题中常用的预处理步骤。至此，通过候选结节检测算法，本文筛选得到了所有疑似结节，包括它们的三维位置、直径、结节属性概率分布。

1.2　基于注意力的多尺度结节分类网络

从前一阶段获得了候选结节后，这一阶段需要从候选结节中准确地分辨出真正的结节。为此，本文设计了一个基于三维卷积神经网络和残差模块的多尺度二分类网络。残差网络错误!未找到引用源。解决了网络深度加深以后的性能退化问题，因而在多个图像任务中表现出强大的性能。在提出的分类网络中加入了残差模块$\boldsymbol{y} = \text{F}(\boldsymbol{x}, \{\boldsymbol{W}_i\}) + \boldsymbol{x}$ ，$\boldsymbol{x}$ 与 $\boldsymbol{y}$ 分别是模块的输入与输出，$\text{F}(\cdot)$是一个包含卷积、batch normalization（BN）错误!未找到引用源。、rectified linear unit（ReLU）的非线性变换，具体结构见图 2。在这个分类任务中加入残差模块将有助于误差反向传播，同时使网络具有更强的表达能力错误!未找到引用源。，从而有助于模型优化。

该阶段常以候选结节的质心为中心提取固定尺寸的三维块并对块进行二分类，所有预测概率大于阈值的候选结节即为检测到的肺结节。一般情况下，分类网络的输入包括固定尺寸的图像与对应类别，这存在两个问题：第一，单一尺度的图像块无法表达不同尺寸的结节信息。正确分类要求三维块中不仅包含结节本身，还应包含结节附近的肺实质环境信息，这对于正确分类也是非常关键的。但由于肺结节的尺寸范围广，对于尺寸较大的结节需要更大的块用于分类，否则可能无法包含所有结节信息，并且只能包含少量的环境上下文信息；对于尺寸较小的结节，如果块太大会导致块中包含的结节信息占比过小而环境信息冗余。第二，仅将图像块作为输入损失了一部分标注信息，包括肺结节在块中的位置、占比，这部分结节信息对于分类是至关重要的。如果对输入图像块的每个体素赋予权重，那么接近结节中心的体素应该获得更大的权值，而距离结节中心越远则影响越小。为了解决上述两个问题，本文设计了基于注意力的多尺度网络结构错误!未找到引用源。。

图像中的各个部分对分类的重要性不完全相同，视觉注意力机制需要找到图像中具有判别力的区域错误!未找到引用源。错误!未找到引用源。，而这里的注意区域是可以直接获得的，注意力中心即为结节质心。具体来说，对于训练阶段，正样本可直接由标注获得结节质心位置与直径，负样本则由候选结节检测网络进行难例筛选后得到，同时可得假阳性结节的质心与直径，对于测试阶段，同样可由候选结节检测网络获得候选结节的三维位置和直径。

给定三维图像块$(\text{X}, \text{Y}, \text{L})$，其中 $\boldsymbol{L}$ 表示输入三维块的真结节或假阳结节信息(x,y,z,d)，网络结构的隐层包含两个平行独立的分支，局部分支以结节质心为中心提取局部注意区域，区域尺寸为结节尺寸的 2 倍，并将其重采样至固定尺寸，全局分支直接以完整图像块作为输入，两个分支的网络结构如图 2，其中全局分支为括号内的配置，其他均与局部分支一致，可以分别得到基于全局块和局部块的二类概率分布$\text{p}(\hat{\text{Y}}|\text{X}, \text{W}_{\text{global}})$和

$p(\hat{Y}|\boldsymbol{X},\boldsymbol{L},\boldsymbol{W}_{loc})$，最终网络输出的二类概率分布通过这两个概率分布的加权组合得到如下：

$$p(\hat{Y}=1|\boldsymbol{X},\boldsymbol{L},\boldsymbol{W}_{\text{global}},\boldsymbol{W}_{\text{loc}}) = \alpha\cdot p(\hat{Y}=1|\boldsymbol{X},\boldsymbol{L},\boldsymbol{W}_{\text{loc}}) + (1-\alpha)\cdot \text{p}(\hat{Y}=1|\boldsymbol{X},\ \boldsymbol{W}_{\text{global}}) \quad (5)$$

权重α是由两个分支的 Softmax 层经过两个全连接得到的隐层，见图 1，因此它也是由网络自学习得到的，可以灵活地根据两个分支鉴别能力的强弱进行加权组合。最后得到分类网络的损失函数如下：

$$L = -\frac{1}{N}\sum_i \log p(Y^i|\boldsymbol{X}^i,\boldsymbol{W}_{\text{global}},\boldsymbol{W}_{\text{loc}}) + \lambda(\|\boldsymbol{W}_{\text{global}}\|_2^2 + \|\boldsymbol{W}_{\text{loc}}\|_2^2) \quad (6)$$

其中：第一项是由对数似然函数定义的结节分类损失；第二项是正则项；为了防止过拟合，λ是权重衰减参数，用于调节模型复杂度对损失函数的影响。

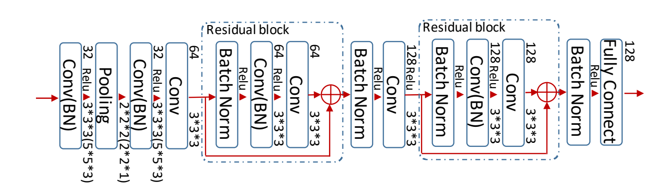


图 2　结节分类网络的局部分支结构

2　实验设计与结果分析

本文在大规模公开数据集 LUNA16[21]上验证算法。该数据集来自于公开数据库 LIDC-IDRI 错误!未找到引用源。，由专家经过一系列筛选和标注后得到。LUNA16 包含 888 张低剂量 CT，同时给出了肺结节中心位置和直径的标注。实验时，将 LUNA16 数据集随机分成 10 个子集进行交叉验证。评价指标包括灵敏度和每张扫描的平均假阳数目（FPs/scan）。LUNA16 竞赛还通过 competition performance metric（CPM）[21]衡量算法性能。CPM 指在不同 FPs/scan（1/8, 1/4, 1/2, 1, 2, 4, 8）下的平均灵敏度。对整个系统而言，当检测到的结节位于真正结节的半径范围内，即认为这是一次正确的检测。

2.1　实现细节

由于 CT 扫描可能来自不同设备，切片厚度不完全相同，所以首先将数据重采样至 1mm*1mm*1mm，保证数据的一致性。为了增加样本的多样性，对正样本进行了数据扩增，包括在 xyz 方向进行[-8, 8]随机平移、翻转，[0.9, 1.1]的随机尺度变化，随机全旋转[-180°, 180°]，设定这些参数能够尽量使结节留在原结节质心附近的 30*30*30 的立方体内。候选结节检测阶段的训练块尺寸为(56*56*56)，结节分类阶段的输入块尺寸为(60*60*30)，局部块的尺寸为（20*20*20）。网络采用随机梯度下降法进行优化，预训练 3D U-net 的权重随机初始化为高斯分布 $N(0,0.01)$，初始学习率设置为 0.001，联合调优网络的初始学习率设置为 0.0001，正负样本比例 1：1。候选结节检测阶段的难例选取阈值与极值检测阈值均设置为 0.8。结节分类网络的参数随机初始化同**错误!未找到引用源。**。结节分类损失中的参数λ设置为 1e-4。本文算法在 Nvidia Titan X GPU 上使用基于 Lasagne 深度学习框架的 Theano 实现。

2.2　实验结果与分析

2.2.1 候选结节检测结果

为了验证提出的候选结节检测算法的有效性，设计了三组对比实验，它们都采用相同的简化 3D U-net 网络结构。首先，本文分别训练了基于交叉熵损失与 Dice 损失的模型，实验结果见表 1。对比这两组结果，与交叉熵损失相比，采用 Dice 损失大大提高了候选结节检测的灵敏度，这说明 Dice 损失更能检出所有疑似结节，但是假阳数有所增加。进一步地，本文加入了难例筛选与联合调优机制，数据集共包含负样本 551 065 例，经过难例筛选后检测到包含假阳结节的负样本 35 118 例，说明有约 6.4%的负样本与结节相似，它们与正样本进行联合调优得到最终的检测模型。从表 1 的结果来看，该机制降低了假阳性结节的误检率，FPs/scan 从 86.9 降低至 59.1，同时保持了高灵敏度，这说明难例联合调优增强了网络对于真假阳性样本的鉴别能力。为了进一步证明难例筛选与预训练模型的必要性，本文设计了一个直接由正负样本基于混合损失训练的网络。从结果可以看出，与本文提出的候选结节检测算法相比，将不经筛选的负样本与正样本直接联合训练得到的网络鉴别能力与灵敏度都有所降低。

表 1　候选结节检测对比算法结果

方法	sensitivity	FPs/scan
基于交叉熵损失	82.86%	62.5
基于 Dice 损失	97.14%	86.9
基于混合损失	95.23%	80.2
提出的算法	**97.18%**	**59.1**

2.2.2 结节分类结果

对于结节分类阶段，为了验证残差模块与多尺度结构的有效性，本文实现了三个网络结构，分别是传统卷积分类网络、加入残差模块的卷积分类网络和基于注意力的多尺度残差分类网络，并采用受试者特征（free-response receiver operating characteristic, FROC）曲线错误!未找到引用源。和 CPM 错误!未找到引用源。来定量分析。FROC 曲线可以直观地反映不同假阳性数目下的系统灵敏度，实验结果见图 3。三个网络结构在结节分类阶段的 CPM 分别为 0.832，0.848 和 0.880，说明在这一任务中，残差模块有助于模型优化并提升网络的性能，而基于注意力的多尺度结构则有效利用了疑似结节的位置和直径信息，进一步提高分类网络的鉴别能力，证明了多尺度结构的有效性。图 3 进一步展示了提出的算法与已有肺结节检测算法[15]**错误!未找到引用源。**的性能比较。由图可见，多尺度残差网络获得了最佳性能。

在提出的多尺度分类算法中，两个分支预测概率的权重是由网络自学习一组权重参数得到的。为了验证自学习权重参数的有效性，设计了一系列对比实验将权重设置为固定值 $\alpha = 0\sim1(\text{step} = 0.1)$，从而得到在不同权重下的多尺度网络

CPM，如

图 4 所示。结果表明固定权重 α 时，α 取较小或较大都只利用了单一分支的信息，性能不佳；而当两个分支的融合比例接近时，更能充分利用两个分支的结果，性能达到最优。进一步地，提出算法的平均灵敏度 0.880 高于任意固定权重的算法，说明提出的算法可以根据两个分支的表征能力自适应地选择不同的权重，从而更有效地利用两个分支信息。

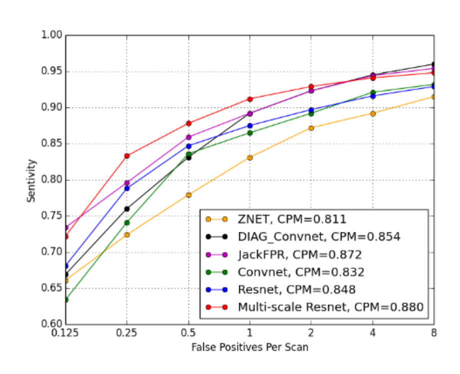


图 3　不同网络、算法的 FROC 曲线及 CPM 比较

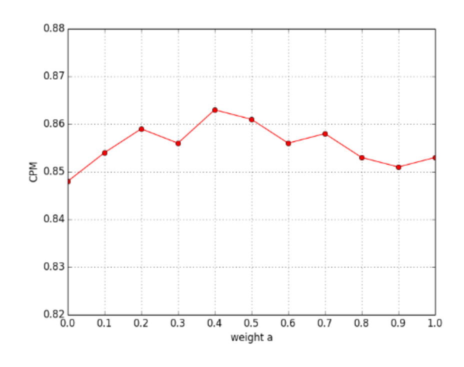


图 4　不同权值的多尺度网络的 CPM 比较

3　结束语

本文针对 CT 图像的肺结节自动检测任务提出了一个改进的方法，该方法首先利用改进 U-net 及难例挖掘并联合调优的网络检测候选结节，它在保证高灵敏度的同时降低了假阳率；然后采用基于注意力的多尺度残差网络对候选结节进行分类，充分利用了结节位置及尺寸信息并解决了结节尺度差异大的问题。LUNA16 上的实验结果验证了改进算法的有效性，同时超过了已有算法的性能。本文方法针对正负样本不平衡与结节尺度差异大这两个问题进行了探索，并获得了更优的结果，这也是医学图像研究中普遍存在的问题。未来会将这些方法应用于更多医学图像领域，并结合具体问题探索更优的方法与策略。

参考文献：

[1] Sluimer I, Schilham A, Prokop M, *et al.* Computer analysis of computed tomography scans of the lung: a survey [J]. IEEE Trans on Medical Imaging, 2006, 25 (4): 385-405.

[2] Messay T, Hardie R C, Rogers S K. A new computationally efficient CAD system for pulmonary nodule detection in CT imagery [J]. Medical Image Analysis, 2010, 14 (3): 390-406.

[3] Tan M, Deklerck R, Jansen B, *et al.* A novel computer-aided lung nodule detection system for CT images [J]. Medical Physics, 2011, 38 (10): 5630-5645.

[4] Setio A A A, Jacobs C, Gelderblom J, *et al.* Automatic detection of large pulmonary solid nodules in thoracic CT images [J]. Medical Physics, 2015, 42 (10): 5642-5653.

[5] Lopez Torres E, Fiorina E, Pennazio F, *et al.* Large scale validation of the M5L lung CAD on heterogeneous CT datasets [J]. Medical Physics, 2015, 42 (4): 1477-89.

[6] Murphy K, Van Ginneken B, Schilham A M, *et al.* A large-scale evaluation of automatic pulmonary nodule detection in chest CT using local image features and k-nearest-neighbour classification [J]. Medical Image Analysis, 2009, 13 (5): 757-770.

[7] Cover T, Hart P. Nearest neighbor pattern classification [J]. IEEE Trans on Information Theory, 1967, 13 (1): 21-27.

[8] Jacobs C, Van Rikxoort E M, Twellmann T, *et al.* Automatic detection of subsolid pulmonary nodules in thoracic computed tomography images [J]. Medical Image Analysis, 2014, 18 (2): 374–384.

[9] Friedman J, Hastie T, Tibshirani R. Additive logistic regression: a statistical view of boosting [J]. Annals of Statistics, 2000, 28 (2): 337-374.

[10] Retico A, Delogu P, Fantacci M E, *et al.* Lung nodule detection in low-dose and thin-slice computed tomography [J]. Computers in Biology & Medicine, 2008, 38 (4): 525-534.

[11] Van Ginneken B, Setio A A A, Jacobs C, *et al.* Off-the-shelf convolutional neural network features for pulmonary nodule detection in computed tomography scans [C]// Proc of the 12th IEEE International Symposium on Biomedical Imaging. Piscataway, NJ: IEEE Press, 2015: 286-289.

[12] Jin Xinyu, Zhang Yuchen, Jin Qiliang. Pulmonary nodule detection based on CT images using convolution neural network [C]// Proc of the 9th IEEE International Symposium on Computational Intelligence and Design. Piscataway, NJ: IEEE Press, 2017: 202-204.

[13] Golan R, Jacob C, Denzinger J. Lung nodule detection in CT images using deep convolutional neural networks [C]// Proc of IEEE International Joint Conference on Neural Networks. Piscataway, NJ: IEEE Press, 2016: 243-250.

[14] Anirudh R, Thiagarajan J, Bremer T, *et al.* Lung nodule detection using 3D convolutional neural networks trained on weakly labeled data [C]// Proc of

SPIE Medical Imaging. 2016: 978532.

[15] Huang Xiaojie, Shan Junjie, Vaidya V. Lung nodule detection in CT using 3D convolutional neural networks [C]// Proc of the 14th IEEE International Symposium on Biomedical Imaging. Piscataway, NJ: IEEE Press, 2017: 379-383.

[16] Setio A A A, Ciompi F, Litjens G, *et al.* Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks [J]. IEEE Trans on Medical Imaging, 2016, 35 (5): 1160-1169.

[17] Dou Di, Chen Hao, Yu Lequan, *et al.* Multilevel contextual 3-D CNNs for false positive reduction in pulmonary nodule detection [J]. IEEE Trans on Bio-medical Engineering, 2017, 64 (7): 1558-1567.

[18] Ren Shaoqing, He Kaiming, Girshick R, *et al.* Faster R-CNN: towards real-time object detection with region proposal networks [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2017, 39 (6): 1137-1149.

[19] Zhu Wentao, Liu Chaochun, Fan Wei, *et al.* DeepLung: deep 3D dual path nets for automated pulmonary nodule detection and classification [C]// Proc of IEEE Workshop on Applications of Computer Vision. Piscataway, NJ: IEEE Press, 2018.

[20] Ding Jia, Li Aoxue, Hu Zhiqiang, *et al.* Accurate pulmonary nodule detection in computed tomography images using deep convolutional neural networks [C]// Proc of International Conference on Medical Computing and Computer-Assisted Intervention. Berlin: Springer, 2017: 559-567.

[21] Sudre C H, Li Wenqi, Vercauteren T, *et al.* Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations [C]// Proc of Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Berlin: Springer, 2017: 240-248.

[22] Setio A A A, Traverso A, De Bel T, *et al.* Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the LUNA16 challenge [J]. Medical Image Analysis, 2017, 42: 1-13.

[23] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation [C]// Proc of International Conference on Medical Computing and Computer-Assisted Intervention. Berlin: Springer, 2015: 234-241.

[24] Shrivastava A, Gupta A, Girshick R. Training region-based object detectors with online hard example mining [C]// Proc of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2016: 761-769.

[25] Van Rikxoort E, De Hoop B, Viergever M A, *et al.* Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection [J]. Medical Physics, 2009, 36 (7): 2934-2947.

[26] He Kaiming, Zhang Xiangyu, Ren Shaoqing, *et al.* Identity mappings in deep residual networks [C]// Proc of European Conference on Computer Vision. Berlin: Springer, 2016: 630-645.

[27] Loffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C]// Proc of International Conference on Machine Learning. New York: ACM Press, 2015: 448-456.

[28] Veit A, Wilber M, Belongie S. Residual networks behave like ensembles of relatively shallow networks [C]// Proc of the 29th International Conference on Neural Information Processing Systems. Cambridge, MA: MIT Press, 2016.

[29] Kamnitsas K, Ledig C, Newcombe V F, *et al.* Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation [J]. Medical Image Analysis, 2016, 36: 61-78.

[30] Xiao Tianjun, Xu Yichong, Yang Kuiyuan, *et al.* The application of two-level attention models in deep convolutional neural network for fine-grained image classification [C]// Proc of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2015: 842–850.

[31] Mnih V, Heess N, Graves A, *et al.* Recurrent models of visual attention [C]// Proc of the 27th International Conference on Neural Information Processing Systems. Cambridge, MA: MIT Press, 2014: 2204–2212.

[32] Armato S G, Mclennan G, Bidaut L, *et al.* The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans [J]. Medical Physics, 2011, 38 (2): 915-931.